JGR Solid Earth

RESEARCH ARTICLE
10.1029/2019JB017874

Improving Physics-Based Aftershock Forecasts During the 2016–2017 Central Italy Earthquake Cascade

S. Mancini[1,2], M. Segou[2], M. J. Werner[1], and C. Cattania[3]

[1]School of Earth Sciences, University of Bristol, Bristol, UK, [2]British Geological Survey, Lyell Centre, Edinburgh, UK, [3]Department of Geophysics, Stanford University, Stanford, CA, USA

Abstract The 2016–2017 Central Apennines earthquake sequence is a recent example of how damages from subsequent aftershocks can exceed those caused by the initial mainshock. Recent studies reveal that physics-based aftershock forecasts present comparable skills to their statistical counterparts, but their performance remains a controversial subject. Here we employ physics-based models that combine the elasto-static stress transfer with rate-and-state friction laws, and short-term statistical Epidemic Type Aftershock Sequence (ETAS) models to describe the spatiotemporal evolution of the earthquake cascade. We then track the absolute and relative model performance using log-likelihood statistics for a 1-year horizon after the 24 August 2016 $M_w = 6.0$ Amatrice earthquake. We perform a series of pseudoprospective experiments by producing seven classes of Coulomb rate-state (CRS) forecasts with gradual increase in data input quality and model complexity. Our goal is to investigate the influence of data quality on the predictive power of physics-based models and to assess the comparative performance of the forecasts in critical time windows, such as the period following the 26 October Visso earthquakes leading to the 30 October $M_w = 6.5$ Norcia mainshock. We find that (1) the spatiotemporal performance of the basic CRS models is poor and progressively improves as more refined data are used, (2) CRS forecasts are about as informative as ETAS when secondary triggering effects from $M3+$ earthquakes are included together with spatially variable slip models, spatially heterogeneous receiver faults, and optimized rate-and-state parameters. After the Visso earthquakes, the more elaborate CRS model outperforms ETAS highlighting the importance of the static stress transfer for operational earthquake forecasting.

Key Points:
- Elaborate physics-based forecasts perform close to short-term statistical models
- The predictive power of physics-based models increases when source and fault heterogeneities are considered
- In the days preceding the $M_w = 6.5$ Norcia earthquake, the near-source seismicity patterns are well explained by static stress transfer

Supporting Information:
- Supporting Information S1

Correspondence to:
S. Mancini,
simone.mancini@bristol.ac.uk

Citation:
Mancini, S., Segou, M., Werner, M. J., & Cattania, C. (2019). Improving physics-based aftershock forecasts during the 2016–2017 Central Italy Earthquake Cascade. *Journal of Geophysical Research: Solid Earth*, *124*. https://doi.org/10.1029/2019JB017874

Received 18 APR 2019
Accepted 24 JUL 2019
Accepted article online 31 JUL 2019

1. Introduction

Earthquake cascades with multiple moderate to large magnitude events over weeks, months, or years expand the damage zone extensively causing disruption to livelihoods. Operational earthquake forecasts seek to provide reliable real-time information about the time dependence of earthquake hazard to the public and to decision makers (Jordan et al., 2011, 2014), but still little is known about earthquake nucleation processes and short-term fault interaction. For the last 30 years, statistical Epidemic Type Aftershock Sequence (ETAS) forecasts (Ogata, 1988, 1998) have shown considerable skills in capturing the clustering characteristics of triggered seismicity. However, these models offer limited insight into the physics of earthquake nucleation and fault interaction in terms of continuum mechanics. A commonly accepted physical mechanism to explain earthquake triggering is the static stress transfer hypothesis (Harris & Simpson, 1992). Nowadays, the implementation of laboratory-derived friction laws describing the seismicity response to an earthquake perturbation (Dieterich, 1994) is the basis of many physics-based forecasts known as Coulomb rate-state (CRS) models (Cattania et al., 2018; Cocco et al., 2010; Parsons et al., 2012, 2014; Segou, 2016; Toda & Enescu, 2011; Toda et al., 2005, among others). Recent advances in CRS modeling demonstrated that addressing the uncertainties behind physics-based models leads to a predictive power comparable to empirical models (Cattania et al., 2018; Segou & Parsons, 2016). However, the limitation for an operational use of CRS models is that their parametrization goes beyond a mere earthquake catalog and requires representations of earthquake sources and nearby receiver faults that are unlikely to be available immediately after a major event.

The Collaboratory for the Study of Earthquake Predictability (CSEP) provides a platform to evaluate the performance of forecast models in retrospective and prospective mode (Jordan, 2006; Michael & Werner, 2018). Here we assess how data quality and individual model choices driven by real-time conditions affect the performance of physics-based forecasts benchmarked against a statistical ETAS model. The 2016–2017

central Apennines (Italy) seismic sequence is an ideal testing ground with nine $M5+$ earthquakes between August 2016 and January 2017 that devastated many villages and historical heritage buildings.

We present seven classes of physics-based models with gradual complexity increase in a pseudoprospective experimental framework for the first year after the $M_w = 6.0$ Amatrice event. The simplest forecasts include preliminary data available a few minutes after each $M_w \geq 5.4$ event, featuring synthetic source models with empirically derived fault length and previously determined fault constitutive parameters. More complex models incorporate (1) optimized rate-state parameters, (2) spatially heterogeneous receiver fault planes, (3) best-available slip models, and (4) secondary triggering effects. We evaluate the absolute and relative model performance using CSEP's metrics over different time horizons.

Our results show that the preliminary assumptions made to fill the early postdisaster knowledge gap severely hamper the absolute performance of CRS forecasts. When we incorporate revised data, optimize the rate-and-state parametrization on previous regional seismicity, and account for the multilevel heterogeneities as (1) 3D spatial distribution of receivers, (2) spatially variable fault slip models, and (3) smaller magnitude aftershocks that reshape the local stress field, we obtain a dramatic performance improvement of physics-based models. Results suggest that CRS performance is comparable to ETAS only when secondary triggering is taken into account.

1.1. The Central Apennines Earthquake Sequence

The Central Apennines are among the highest seismic hazard zones in Europe (Woessner et al., 2015). Present day deformation is taken up by NW-SE trending normal fault systems, expressing the postorogenic extension at a rate of ~3 mm/year (Serpelloni et al., 2005). Historical and modern seismicity indicate moderate (M5+) to large (M6+) magnitude earthquake cascades associated with heavy damage patterns in the broader area (Rovida et al., 2016) such as the 1703 Norcia cascade where three $M6.2+$ earthquakes occurred within less than 20 days (Boschi et al., 2000), the six $M5.2+$ events that struck Colfiorito between September and October 1997 (Amato et al., 1998; Chiaraluce et al., 2003, 2004), and the 2009 $M_w = 6.3$ L'Aquila sequence with an accelerating seismicity pattern observed few days before the mainshock (Chiarabba et al., 2009; Chiaraluce et al., 2011).

The 24 August 2016 Amatrice earthquake activated a 60-km long normal fault system (Figure 1a) and was followed by an $M_w = 5.4$ aftershock within less than 1 hr at the northern updip part of the mainshock rupture (Chiaraluce et al., 2017, and references therein). Two months later, on 26 October, two $M_w = 5.4$ and $M_w = 5.9$ earthquakes occurred further north near the village of Visso, closer to the southernmost aftershock zone of the 1997 Colfiorito sequence. The 30 October $M_w = 6.5$ Norcia event remains the largest event of the sequence to date and the strongest in Italy since the 1980 $M_w = 6.9$ Irpinia earthquake. Few months later, on 18 January 2017, four $M5+$ earthquakes occurred within a 4-hr window to the south of the Norcia mainshock rupture, coinciding with the northern aftershock zone of the 2009 L'Aquila sequence. The January 2017 seismicity raised immediate concerns about the Campotosto artificial dam lake, the second largest man-made lake in Europe.

In Table 1 we summarize the preliminary and revised source parameters of the major earthquakes of the Amatrice-Visso-Norcia (hereinafter, AVN) sequence.

2. Data

We implement our models within a 3D grid (0–12 km of depth) with 2-km spacing in a ~150 × 150 km testing region (Figures S1 and S2 in the supporting information). We use (1) the seismicity catalog of the Italian Seismological Instrumental and Parametric Database for the period between 1 January 1990 and 23 August 2016 (learning phase, 1533 $M3+$ events), (2) the near-real-time catalog for the 1 year after the $M_w = 6.0$ Amatrice earthquake (testing phase, 1160 $M3+$ events), and (3) centroid-moment-tensor (CMT) catalogs for the learning phase (Figure S1) and the Database of Individual Seismogenic Sources (DISS, Basili et al., 2008; DISS Working Group, 2018) for constraining the active seismogenic sources and regional rupture styles of the central Apennines.

For the primary events ($M_w \geq 5.4$), we estimate the coseismic stress changes using the finite-fault slip models of Tinti et al. (2016) and Scognamiglio et al. (2016) for the Amatrice I–II events, respectively, and Chiaraluce

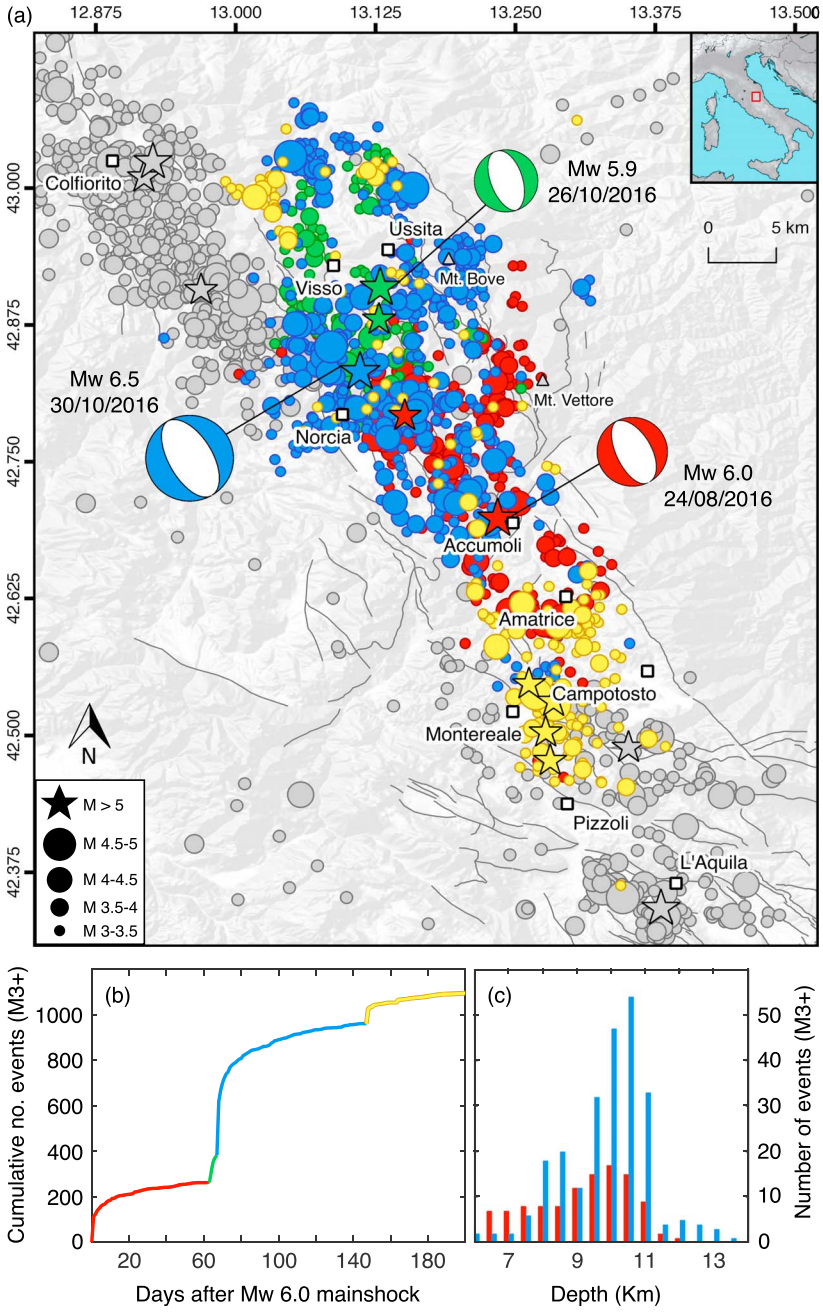

Figure 1. The 2016–2017 AVN sequence (M3+). (a) Map view. Colors indicate different seismicity phases: red, 24 August 2016 to 25 October 2016; green, 26 October to 29 October 2016; blue, 30 October 2016 to 17 January 2017; yellow, 18 January to 24 August 2017. The focal mechanisms of the three largest events are also displayed. Gray symbols indicate the 1997 Colfiorito and 2009 L'Aquila seismic sequences (M3+). We report the mapped active faults of the region (EMERGEO Working Group, 2016), (b) cumulative seismicity (M3+) with time, and (c) depth distribution of M3+ events in the first 24 hr following the M_w = 6.0 Amatrice (red bars) and M_w = 6.5 Norcia (blue bars) mainshocks. AVN = Amatrice-Visso-Norcia.

et al. (2017) for the Visso II and Norcia events. For the 18 January 2017 M_w = 5.5 and M_w = 5.4 Campotosto earthquakes, we use the only preliminary slip models available, as no refined versions have been issued. For the sources that lack a rupture model, we construct a synthetic slip distribution from their moment tensor solution, with empirically derived fault dimensions (Wells & Coppersmith, 1994) and uniform slip (from the moment relation of Hanks & Kanamori, 1979) tapered at the edges.

Table 1
Source Parameters for the M5+ Events of the AVN Sequence

			Preliminary data		Revised data		
Event	Date	Time (UTC)	M_L	Z (km)	M_w	Z (km)	Slip model
Amatrice I	24/08/2016	01:36:32	6.0	4.0	6.0	8.0	Yes
Amatrice II	24/08/2016	02:46:36	5.4	9.0	5.4	8.0	Yes
Visso I	26/10/2016	17:20:02	5.4	9.0	5.4	9.0	No
Visso II	26/10/2016	19:27:30	5.9	8.0	5.9	8.0	Yes
Norcia	30/10/2016	06:58:20	6.1	10.0	6.5	9.0	Yes
Campotosto I	18/01/2017	09:25:40	5.3	9.0	5.1	9.0	No
Campotosto II	18/01/2017	10:14:09	5.4	9.0	5.5	11.0	Yes
Campotosto III	18/01/2017	10:25:23	5.3	9.0	5.4	11.0	Yes
Campotosto IV	18/01/2017	13:33:36	5.1	10.0	5.0	10.0	No

Note. Time stamps and magnitudes are available at http://cnt.rm.ingv.it, while the revised hypocentral depths are taken from the corresponding slip models where available. The period between the moment an event is recorded and the time when its rupture model is estimated varies from few days to several weeks. For the AVN sequence, finite-length slip models were computed for six out of the nine *M*5+ events, and for most of them a period of 2–3 weeks from their occurrence passed before the release of a first robust version. Dates are formatted as day/month/year. AVN = Amatrice-Visso-Norcia.

3. Methods

3.1. CRS Modeling

We estimate the coseismic Coulomb stress change (ΔCFF) by using the formulation of Rice (1992):

$$\Delta CFF = \Delta\tau + \mu'(\Delta\sigma_n), \tag{1}$$

where $\Delta\tau$ is the shear stress change on a given receiver fault plane (positive in direction of fault slip), $\Delta\sigma_n$ is the normal stress change, $\mu' = \mu(1 - B_k)$ is the effective coefficient of friction, and B_k is Skempton's coefficient that accounts for pore fluid pressure effects. We set $\mu' = 0.4$ for all models, although we also tested the effect of implementing a lower μ' (see Text S1.3). We assume an elastic half-space with a shear modulus of 30 GPa and Poisson's ratio $\nu = 0.25$.

The receiver planes formulation follows two approaches: the *optimally oriented planes* (King et al., 1994) that assumes that earthquakes nucleate on hypothetical faults in favorable orientation with respect to the regional stress field and the geological receiver plane (GRP) approach. Both approaches received extensive criticism in recent years: the optimally oriented planes for relying on the knowledge of the largely unknown regional stress tensor to resolve stress changes on hypothetical planes that might not even exist (Segou & Parsons, 2016), whereas GRP may miss unmapped faults, when even in well-studied regions such as California ~30% of seismicity occurs on previously unidentified structures (Field et al., 2014). Here we adopt the Segou et al. (2013) approach where GRP is informed by presequence focal mechanisms and active fault maps to achieve a representation of the spatial structural heterogeneity.

We use the coseismic stress change estimates within a rate-and-state friction framework (Dieterich, 1994) to derive the time-dependent evolution of seismicity, expressed in expected number of events within 1-day time intervals for a 1-year time horizon. The seismicity rate R is related to the background rate r_0:

$$R = \frac{r_0}{\gamma\dot{\tau}}, \tag{2}$$

where $\dot{\tau}$ represents the secular shear stressing rate and γ is a state variable. Under a constant secular shear stressing rate, γ reaches the steady state with a value given by

$$\gamma_0 = \frac{1}{\dot{\tau}}. \tag{3}$$

In the absence of stress perturbations, the seismicity rate R equals the background rate r_0. At the time of a stress step the state variable value is given by

$$\gamma_{\mathrm{n}} = \gamma_{\mathrm{n-1}} \exp\left(\frac{-\Delta \mathrm{CFF}}{A\sigma}\right), \tag{4}$$

where ΔCFF is the coseismic stress change, and $A\sigma$ is a constitutive parameter accounting for the effective normal stress acting on the fault. A stress increase causes a drop in γ, resulting in a higher seismicity rate through equation (2). Due to the transient nature of the seismicity rate change, the evolving values of γ produce an Omori decay and R recovers to the background value in a time equal to t_{a} (Dieterich, 1994):

$$t_{\mathrm{a}} = \frac{A\sigma}{\dot{\tau}}. \tag{5}$$

Given this inverse correlation between the stressing rate and the aftershock duration, it is evident that the seismicity rate on the most slowly stressed faults takes more time to decay toward background values (Stein & Liu, 2009).

To address the influence of the rate-and-state parameters $A\sigma$ and $\dot{\tau}$ in the CRS forecasts, we test (1) predetermined values based on the study by Catalli et al. (2008) of the 1997 Colfiorito sequence and (2) parameters derived from a log-likelihood optimization on pre-AVN seismicity (see Text S1.1) using a common learning phase catalog for all CRS models.

The implementation of CRS models is based on a parallel computer code for calculating seismicity induced by time-dependent stress changes (Cattania & Khalid, 2016).

3.2. The ETAS Model

Statistical ETAS models are often used to describe short-term earthquake clustering (Ogata, 1988; Ogata, 1998). As a benchmark for the CRS forecasts, we use one standard version of ETAS (Seif et al., 2016). Although the focus of our study is on the improvements of CRS models, we acknowledge that other ETAS parametrizations may perform differently. Previous experiments show a good performance of ETAS both in retrospective (Cattania et al., 2018; Helmstetter et al., 2006; Marzocchi, Murru, et al., 2012; Werner et al., 2011; Woessner et al., 2011) and during unfolding sequences (e.g., Marzocchi et al., 2017; Marzocchi & Lombardi, 2009), but some weaknesses related with early catalog incompleteness (Omi et al., 2016; Segou & Parsons, 2016) and the absence of fault interaction effects (e.g., Marzocchi & Lombardi, 2009) have been reported.

The ETAS seismicity corresponds to a point process with a stochastic spatiotemporal branching evolution, where each earthquake triggers its own offspring events, whose number depends on the parent's magnitude and follow an Omori law decay. In the ETAS model, triggered earthquakes can have a larger magnitude than their parent event. The total space-time seismicity rate λ (or *conditional intensity*) is defined as

$$\lambda(x, y, t \mid H_{\mathrm{t}}) = \mu(x, y) + \sum_{\mathrm{i}:t_{\mathrm{i}}<t} g(t-t_{\mathrm{i}}, x-x_{\mathrm{i}}, \mathrm{y}-y_{\mathrm{i}}; M_{\mathrm{i}}) \ , \tag{6}$$

where $\mu(x,y)$ is the background rate, a time-independent, and spatially heterogeneous Poisson process, while the summation term represents the triggering history (H_{t}) from all preceding earthquakes occurring at $t_{\mathrm{i}} < t$. The triggering function is expressed by empirical relations, according to the form of Ogata (1998):

$$g(t, x, y; M) = K_0 e^{\alpha(M-M_{\mathrm{cut}})} \cdot c^{p-1} (t+c)^{-p} (p-1) \cdot f(x, y|M), \tag{7}$$

with normalized temporal and spatial distributions as second and third terms on the right-hand side, respectively. The parameter K_0 regulates the short-term aftershock productivity by a parent event with magnitude M equal or above a minimum triggering magnitude (M_{cut}), here set to 3.0; α establishes the efficiency of earthquakes in triggering aftershocks as a function of magnitude. The second term on the right-hand side of equation (7) is the modified Omori law (Utsu, 1961) describing the temporal distribution of triggered earthquakes. The term $f(x,y|M)$ represents the probability distribution function of the spatial decay of triggered seismicity near the parent event given the parent's magnitude. Here we adopt a spatially isotropic power law distribution (e.g., Ogata & Zhuang, 2006; Werner et al., 2011):

$$f(x,y|M) = \left(d\, e^{\gamma(M-M_{\text{cut}})}\right)^{q-1}/\pi\cdot\left(x^2+y^2+d\cdot e^{\gamma(M_i-M_{\text{cut}})}\right)^{-q}(q-1), \quad (8)$$

where q describes how triggered events decay in space, and the term $d\cdot e^{\gamma(M_i-M_{\text{cut}})}$ scales the size of the M_{cut} aftershock zone with the magnitude of the parent earthquake.

To estimate the ETAS parameters, we use the Maximum Likelihood Estimation method (see Text S2.1) during the ETAS learning phase corresponding to the latter part of the Italian Seismological Instrumental and Parametric Database catalog (2005–2016) after the Italian permanent seismic network was considerably improved (Schorlemmer et al., 2010).

3.3. Development of Forecast Models

In this section, we describe forecast characteristics and model parameters. All models share a 1-year forecast horizon, with 24-hr time windows (dt) and target seismicity magnitude $M3+$.

The reference model CRS-1 is based on (1) real-time preliminary catalog data, including hypocenters and focal mechanisms available within few minutes to 1 hr, (2) stress perturbations from $M_w \geq 5.4$ events estimated using a uniform slip model with kinematic parameters from focal mechanisms with nodal planes selection constrained by the predominant rupture geometries reported in the DISS database (DISS Working Group, 2018), (3) a spatially homogeneous background rate obtained by stochastic declustering (Zhuang et al., 2002) of the CRS learning phase catalog and averaged over the testing region (r_0 = 0.034 M3+ events/day), and (4) spatially uniform receiver planes for ΔCFF estimation expressing the NW-SE striking, SW dipping main trend of the central Apennines normal fault systems (Basili et al., 2008). This latter assumption is valid when we do not know enough about the complex structures of the neighboring faults (McCloskey et al., 2003). This preliminary forecast realization features rate-and-state parameters $A\sigma$ = 0.04 MPa and $\dot{\tau} = 10^{-3}$ MPa/year (t_a=40 years), previously used to characterize the active faults of the 1997 Colfiorito sequence (Catalli et al., 2008). The only difference between models CRS-1 and CRS-2 is that the latter implements a heterogeneous background rate, smoothed in space according to the adaptive kernel method proposed by Helmstetter et al. (2007).

From CRS-3 onward we use (1) revised hypocentral locations and moment tensor solutions available within 1–3 hr after a large earthquake (Table 1), and (2) fault constitutive parameters optimized on the learning phase catalog (see Text S1.1). During the optimization procedure, the grid search for $A\sigma$ ranges between [0.01–0.1] MPa enveloping all previous estimates (e.g., Catalli et al., 2008; Console et al., 2006; Toda et al., 1998) with aftershock durations (t_a) of [1–1000] years. The best fit Aσ-$\dot{\tau}$ couples for each model are shown in Table 2. Otherwise, CRS-3 features the same implementation parameters as model CRS-2.

In model CRS-4, we introduce spatially variable receiver planes (SVP) to achieve a better representation of the structural heterogeneity of the fault system. We start by randomly selecting a nodal plane for each of the focal mechanisms included in the learning phase catalog (Figure S1), and then we assign a rupture plane to each grid point through a 3D nearest neighbor spatial association. In regions where no previous focal mechanism exists, the assignment of a receiver fault is aided by the DISS database of active seismogenic structures (DISS Working Group, 2018). The resulting discrete fault grid is shown in Figure S2.

In CRS-5, we include the finite-length rupture models for events with $M_w \geq 5.4$ to implement a representation of spatial slip variability together with the structural heterogeneity expressed by the SVP receiver matrix in CRS-4.

The CRS-6 and CRS-7 models incorporate secondary triggering effects of smaller magnitudes with different thresholds (CRS-6, 41 $M4+$ events; CRS-7, 1167 $M3+$ events). In CRS-6, we represent 35 $M4+$ events using uniform slip distributions (similar to CRS-1 to CRS-4) with random selection of rupture planes from Time Domain Moment Tensor solutions of the Italian CMT database and peer-reviewed slip models for six events with $M_w \geq 5.4$. In CRS-7, we have rupture characterizations for only 4% of the $M < 4.0$ earthquakes. Therefore, for those events we assume for an isotropic distribution of coseismic stress changes (Helmstetter et al., 2005; Marsan, 2005), following the formulation of Chen et al. (2013):

Table 2
Main Features of CRS Models

		Stress calculations					Rate-and-state parameters			
Model number	Seismic parameters	M_{min}	Secondary Triggering	Slip distribution	μ'	Receivers	r_0	$A\sigma$ (MPa)	$\dot{\tau}$ (MPa/year)	LL optimization
CRS-1	Preliminary	5.4	No	USD	0.4	SUP	Ho	0.04	0.001	No
CRS-2	Preliminary	5.4	No	USD	0.4	SUP	He	0.04	0.001	No
CRS-3	Revised	5.4	No	USD	0.4	SUP	He	0.03	0.00016	Yes
CRS-4	Revised	5.4	No	USD	0.4	SVP	He	0.02	0.00018	Yes
CRS-5	Revised	5.4	No	FLM	0.4	SVP	He	0.02	0.00018	Yes
CRS-6	Revised	4.0	Yes	FLM ($M \geq 5.4$) USD ($M \geq 4.0$)	0.4	SVP	He	0.03	0.00014	Yes
CRS-6s	Revised	4.0	Yes	FLM ($M \geq 5.4$) USD ($M \geq 4.0$)	0.4	SVP	He	0.02	0.00018	No
CRS-7	Revised	3.0	Yes	FLM ($M \geq 5.4$) USD ($M \geq 4.0$) I ($M \geq 3.0$)	0.4	SVP	He	0.015	0.00019	Yes
CRS-7s	Revised	3.0	Yes	FLM ($M \geq 5.4$) USD ($M \geq 4.0$) I ($M \geq 3.0$)	0.4	SVP	He	0.02	0.00018	No

Note. M_{min} = minimum magnitude for stress sources; USD = uniform slip distribution; FLM = finite-length rupture model; I = isotropic stress field; SUP = spatially uniform receiver planes; SVP = spatially variable planes; Ho = homogeneous; He = heterogeneous.

$$\Delta CFF = \frac{M_0}{6\pi r^3}, \tag{9}$$

with M_0 the seismic moment and r the hypocentral distance.

We also produce two additional models, CRS-6s and CRS-7s, that isolate the sole contribution of secondary triggering ($M4+$ and $M3+$ stress sources, respectively) on model performance, where we omit the rate-and-state optimization by implementing the same fault constitutive parameters previously derived for CRS-5. We take these two auxiliary models into consideration during the model validation stage.

In the ETAS model, we set $\alpha = \beta = b \cdot \log(10)$ with Gutenberg-Richter b value = 1, requiring larger magnitude earthquakes to have a higher triggering potential than the small ones. For the estimation of the ETAS parameters, shown in Table S1, we use the $M3+$ ETAS learning phase seismicity within a polygon covering the entire Italian mainland (see Figure S3a and Text S2.1). To account for earthquake interactions outside the spatiotemporal boundaries of the inversion, we also use events in auxiliary spatial and temporal windows (Figure S3). The inverted parameters are in close agreement with those of Seif et al. (2016) for the same areal extent, time window, and M_{cut}. We fix the ETAS parameters for the whole forecast horizon, and we use them to simulate 1,000 synthetic catalogs within each forecast window (dt). Details on the simulation procedure are provided in the supporting information (see Text S2.2).

4. Results

In this section, we present the forecast results in the form of (a) temporal evolution of expected seismicity, (b) maps of predicted earthquake occurrence for period-specific (between mainshocks), and long-term (1 year) time windows. We then focus on the performance evaluation using the N, S, T–test metrics (Rhoades et al., 2011; Zechar et al., 2010) and comment on the absolute and relative predictive power of the models.

4.1. Forecast Time Series

Figure 2a compares the expected and observed daily rates for $M3+$ events. The preliminary CRS-1 and CRS-2 models systematically underestimate the daily seismicity rates by an order of magnitude over the entire period, whereas the first updated model CRS-3 presents a tenfold increase of expected rates arising from the optimized values of the fault constitutive variables.

Refined source parameters together with the optimized rate-and-state variables (CRS-3), the systematic introduction of SVP receiver faults (CRS-4), and heterogeneity in slip models (CRS-5) bring stress-based models closer to the observed rates, especially during the first day following each mainshock.

CRS-5 and CRS-6 share very similar expected rates in the 24 hr following each large magnitude event, standing out among all physics-based models for fitting more closely the short-term (1 day) seismicity

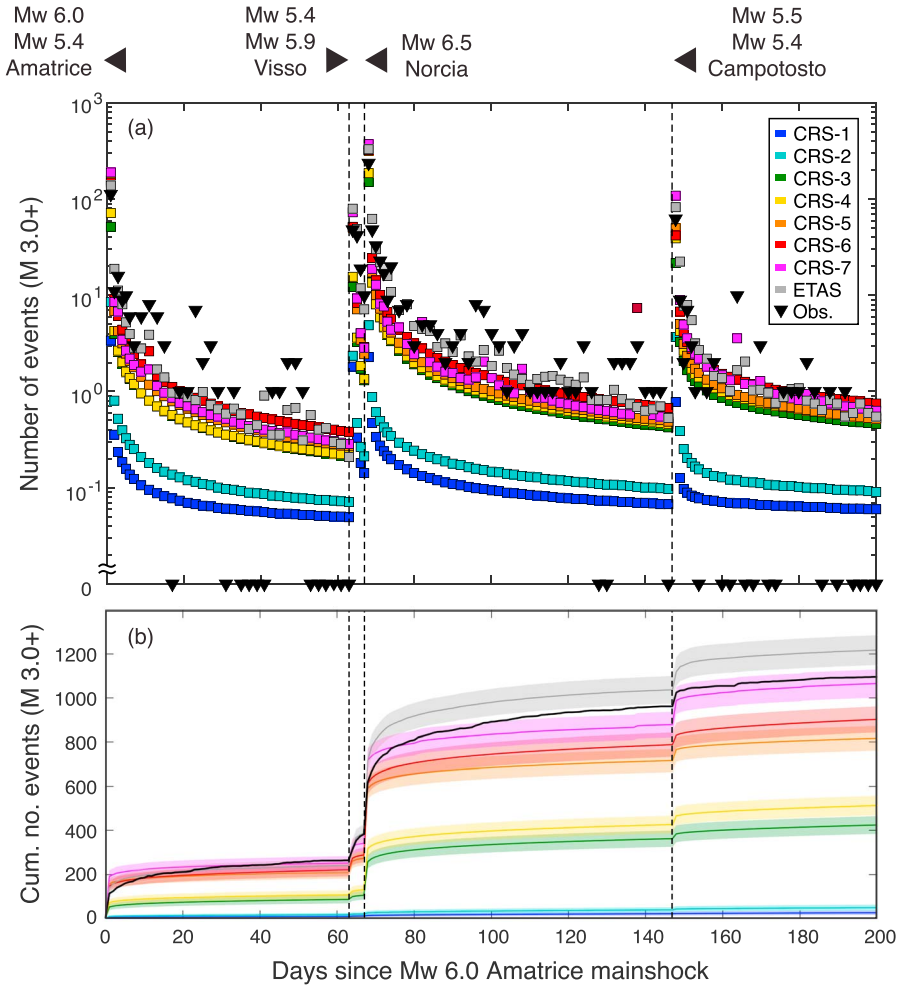

Figure 2. CRS and ETAS forecast time series. (a) Observed (triangles) and forecasted (squares) incremental number of $M3+$ events for 200 days following the M_w = 6.0 Amatrice mainshock. For illustration purposes we plot values at 1-day intervals for each week after the primary events (vertical dashed lines), otherwise at 2-day intervals. (b) Cumulative expected and observed (black solid line) seismicity with shaded areas representing Poissonian uncertainties. CRS = Coulomb rate-state; ETAS = Epidemic Type Aftershock Sequence.

after the M_w = 6.0 Amatrice, M_w = 5.4 Visso, and Campotosto earthquakes. These two model implementations share the same behavior as ETAS after the Amatrice and Norcia mainshocks. CRS-7, with $M3+$ secondary triggering effects, overestimates the short-term ($\leq$ 1 day) seismicity rates after the larger magnitude events (e.g., 376 $M3+$ expected events against 237 observations in the 24 hr following the M_w = 6.5 mainshock); however, this result is biased by early incompleteness of the real-time catalog, as the Mc is larger than 3.0 for at least 6 hr after the Norcia event (Figure S5). For t > 10 days, CRS-6/7 and ETAS closely match the observed rates whereas CRS-3/4/5 generally underestimates.

In Figure 2b we show the cumulative number of expected events versus the cumulative real-time observations. We see that (1) ETAS captures the seismicity decay of the AVN sequence, while CRS models show a faster decay within the first few months, (2) CRS-7 is the only model that adequately forecasts the total number of observations within Poissonian error, and (3) ETAS and the stress-based models from CRS-3 onward satisfactorily estimate the seismicity after the Campotosto earthquakes.

4.2. Seismicity Rate Maps

Figure 3 presents the forecast maps illustrating the difference in spatial distribution of expected rates between the preliminary (CRS-1) and the more elaborate physics-based models (CRS-7) against the statistical ETAS

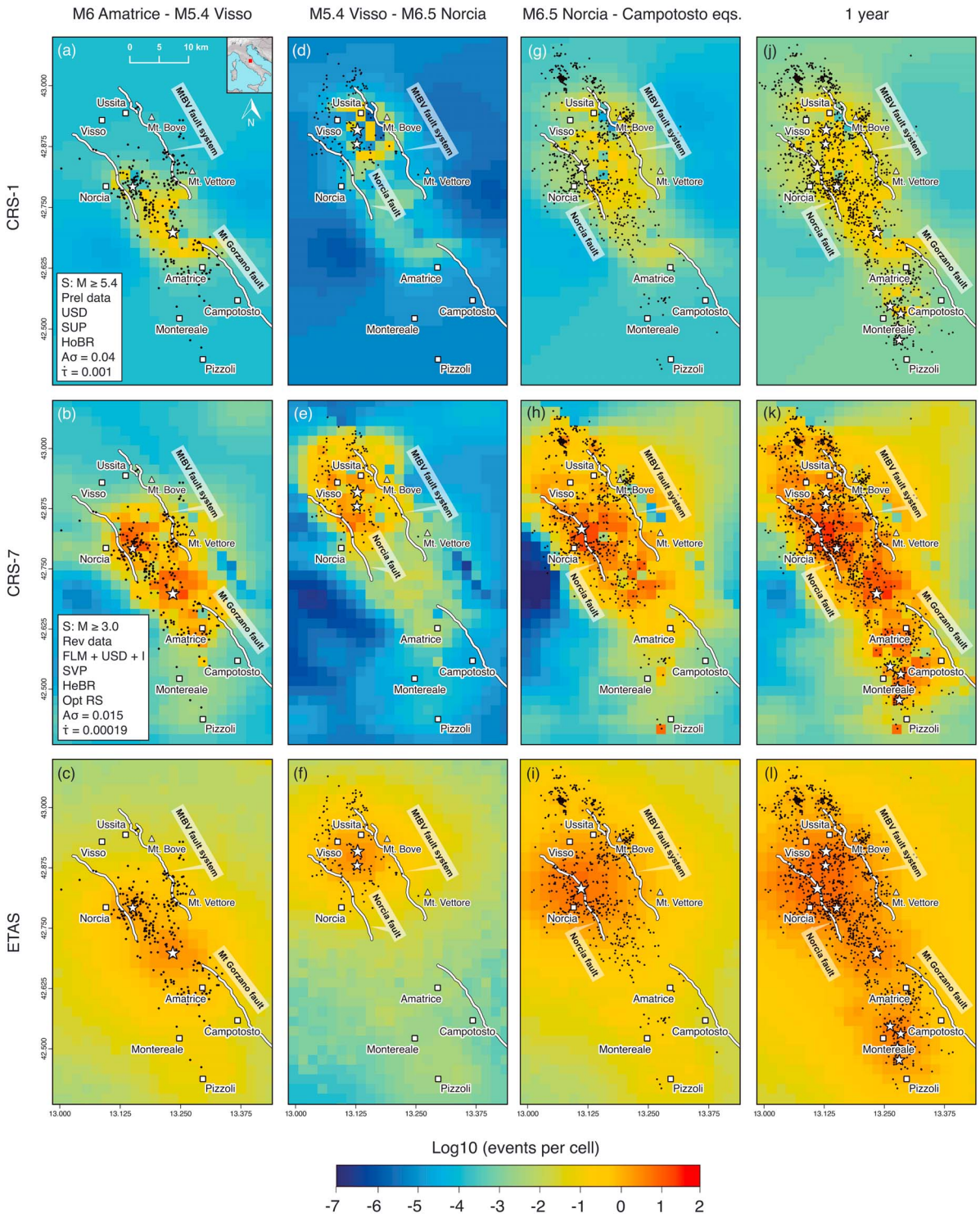

Figure 3. Maps of expected seismicity rates ($M3+$) for CRS-1/7 and ETAS at period-specific windows. We adopt a grid size of 2 × 2 km in all models. Black circles indicate observed events ($M3+$) within each specified time period, while stars indicate the primary earthquakes (not included among the target events of the respective time window). S = sources; Prel = preliminary; Rev = revised; Opt RS = optimized rate-state parameters, USD = uniform slip distribution; SUP = spatially uniform receiver planes; SVP = spatially variable planes; FLM = finite-length rupture model; I = isotropic stress field; HoBR = spatially homogeneous background rate; HeBR = spatially heterogeneous background rate; MtBV = Mount Bove-Mount Vettore. $A\sigma$ values are in MPa, $\dot{\tau}$ values are in MPa/year. CRS = Coulomb rate-state; ETAS = Epidemic Type Aftershock Sequence.

forecast. A full set of forecast maps, including first-day forecasts, is provided in the supporting information (Figures S8–S15).

4.2.1. Amatrice to Visso

The preliminary CRS-1 forecast presents low rates in the near-source region at the northern section of the Mount Gorzano fault, as well as in the epicentral area of the $M_w = 5.4$ aftershock occurred 1 hr after the $M_w = 6.0$ mainshock (Figure 3a). The most advanced CRS-7 model includes SVP receivers, a spatially variable slip model for the Amatrice I/II events, and secondary triggering effects from $M3+$ earthquakes; we observe good visual correlation between the observed and the expected seismicity rates over the entire region with notable increased rates near the two epicenters and close to Amatrice and Mount Vettore (Figure 3b). The ETAS model (Figure 3c) presents a smoother, more isotropic distribution of the higher expected rates over the entire epicentral area in comparison with CRS-7, including locations where no aftershock occurred in the first 2 months.

Comparing between CRS-1 and CRS-3 (Figures S8a–S8c and S9a–S9c), we find that optimized rate-and-state parameters introduce 1–2 orders of magnitude rate increase around the $M_w = 6.0$ event. Comparing CRS-4 to CRS-5 (Figures S8d and S8e and S9d and S9e) shows that heterogeneous slip representations for the Amatrice I/II mainshocks lead to high expected rates in epicentral regions and toward Mount Vettore. In CRS-6 (Figures S8e and S8f and S9e and S9f), we observe that secondary triggering effects from the 17 $M4+$ aftershocks do not exert an obvious difference in the spatial forecast from model CRS-5.

4.2.2. Visso to Norcia

Similar to the previous time window, CRS-1 predicts particularly low rates immediately after and in the region of the $M_w = 5.4$ (Visso I) and $M_w = 5.9$ (Visso II) October events and fails to capture the aftershock activity north of Ussita (Figure 3d). In CRS-7 we observe a noteworthy increase of expected rates in near-source up to 5 orders of magnitudes with respect to CRS-1, and aftershock production is favored on the northern tip of the Mount Bove-Mount Vettore (MtBV) fault system thanks to the spatially variable rupture model for the Visso II event and the inclusion of $M3+$ secondary triggering effects (Figure 3e). The ETAS model presents a good fit with the observed seismicity; the area of higher expected rates extends eastward to Mount Bove, while the area of Norcia presents at least 10 times lower predicted rates compared to the Visso I/II epicentral region (Figure 3f).

4.2.3. Norcia to Campotosto

The preliminary CRS-1, in which stress perturbations are based on simplified uniform rupture distributions, fails to forecast the dramatic seismicity increase following the $M_w = 6.5$ Norcia mainshock, presenting low rates in the near-source region and seismicity suppression on the Mount Bove fault (Figure 3g). The incorporation of the complex Norcia mainshock rupture process, which involved a ~30 km long fault plane, produces (1) near-source coseismic stress heterogeneities and (2) a halving of the stress shadows SW of Mount Vettore and between Ussita and Mount Bove (see Figure S6); this results in a systematic increase of the expected seismicity in the respective areas from CRS-5 onward. The strength of the complex physics-based model CRS-7 is that it better captures the triggered seismicity south of Amatrice due to the incorporation of $M3+$ secondary triggering effects (Figure 3h), while $M4+$ effects in CRS-6 (Figure S12f) present an expected seismicity pattern similar to CRS-5 (Figure S12e). From visual comparison, we find the main differences between ETAS and CRS-7 in the high clustering region around Mount Bove, where the former expects more than 1 order of magnitude higher rates than its stress-based counterpart, and in the near epicentral area, where the aftershock rates predicted by the ETAS model are approximately 2–3 times lower than CRS-7 (Figure 3i). This latter observation is possibly due to the early catalog incompleteness following the Norcia mainshock (Figure S5).

4.2.4. One-Year Forecast

The preliminary CRS-1 (Figure 3j) widely underestimates the observed aftershock production; we see an acceptable spatial distribution of the expected seismicity, but the estimated rates are systematically <1 event/cell over the entire testing region. Model CRS-7 (Figure 3k) addresses this underestimation and is characterized by higher rates, particularly at intermediate off-fault distances (5–10 km from the mainshock faults) where a significant amount of smaller magnitude aftershocks occurred (~45%), contrary to the $M4+$ events that were mostly located at shorter ranges from the main ruptures (Figure S7). The 1-year ETAS forecast (Figure 3l) captures the spatial distribution of the aftershock zone accurately, but due to its branching nature and the lack of a seismicity suppression mechanism, it projects a larger aftershock zone.

Table 3
Short-Term (24 hr After the Primary Events) and Long-Term (1 year) Model Performance

	Short term (24 hr)												Long term (1 year)		
	After Amatrice I			After Visso I			After Norcia			After Campotosto					
Model	jLL_S	$N_{F/O}$	$G_{CRS\text{-}1}$	jLL_S	$N_{F/O}$	$G_{CRS\text{-}1}$	jLL_S	$N_{F/O}$	$G_{CRS\text{-}1}$	jLL_S	$N_{F/O}$	$G_{CRS\text{-}1}$	jLL_S	$N_{F/O}$	$G_{CRS\text{-}1}$
CRS-1	−670	0.03	/	−396	0.04	/	−989	0.01	/	−386	0.01	/	−7104	0.03	/
CRS-2	−665	0.07	0.90	−401	0.05	0.15	−1035	0.02	0.56	−378	0.06	1.65	−7170	0.05	0.58
CRS-3	−760	0.46	1.49	−382	0.26	1.98	−1865	0.64	−0.12	−381	0.34	3.09	−8721	0.41	1.11
CRS-4	−619	0.64	2.89	−303	0.32	3.80	−1614	0.79	1.00	−305	0.62	4.60	−7171	0.49	2.55
CRS-5	−557	1.40	3.62	−130	1.05	7.86	−770	1.34	4.43	−481	0.79	1.88	−6149	0.75	3.47
CRS-6	−673	1.31	2.46	−188	1.08	6.65	−970	1.37	3.69	−422	0.67	2.76	−6771	0.82	3.01
CRS-7	−418	1.69	4.60	−110	1.53	8.17	−681	1.59	4.84	−249	1.73	5.40	−4911	0.97	4.57
ETAS	−222	1.22	6.51	−118	1.65	7.96	−408	1.39	6.06	−112	1.31	7.72	−4326	1.10	5.13

Note. jLL_S = *S*–test joint log-likelihood; $N_{F/O}$ = ratio between forecasted (F) and observed (O) number of events; $G_{CRS\text{-}1}$ = information gain from the preliminary CRS-1 model. Log-likelihood values are negative by definition, and smaller absolute values indicate a better model performance. We note how CRS-3 performance is severely penalized in the 24-hr period after the Norcia mainshock by its poor spatial consistency, leading to a deterioration of the information gain on CRS-1 to a negative value.

4.3. Model Performance in the Testing Region

We evaluate the absolute and relative predictive skills of the forecasts for a 1-year time period. We seek to test (1) the forecasted versus observed number of *M*3+ events using the modified *N*–test scores (Zechar et al., 2010) and rejection ratios (R_N), (2) the spatial consistency of the models through their *S*–test joint log-likelihood scores (jLLs) at short and long term (Schorlemmer et al., 2007; Zechar et al., 2010), and (3) the information gain per earthquake (G, *T*–test; Rhoades et al., 2011) with respect to the preliminary CRS-1 model. A thorough description of the statistical metrics used in this paper is available in the supporting information (see Text S3). In Table 3 we summarize model performance.

Figure 4 presents the 1-day incremental *N*–test scores for the most preliminary (CRS-1) and updated (CRS-7) physics-based models and for the ETAS forecast, where quantiles δ_1 or δ_2 lower than the 0.025 significance level indicate model rejection due to rate underestimation or overestimation, respectively. The incremental *N*–test results for the full set of models are provided as supporting information (Figure S16). We observe that (1) CRS-1 is rejected due to underestimation of the observed number of events ($\delta_1 = 0$) at short term (1 day) and at $t > 1$ week after each mainshock (Figure 4a); (2) CRS-7 and ETAS present ~50% overestimation of the number of events immediately after the Amatrice, Visso, Norcia, and Campotosto events (Figures 4b and 4c); (3) the ETAS model features the lowest R_N (6%) and while it shows a good performance in the time period following the Norcia mainshock, it suffers rejection for 3 days following the Campotosto events. On the other hand, CRS-7 presents a mixed performance after the Amatrice, Visso, and Norcia events, with overestimation within the first 24 hr and underestimation for $t \leq 1$ week, but it passes the test for $t > 24$ hr after the Campotosto seismicity.

CRS-2 performance is similar to CRS-1, as these two preliminary models feature the highest *N*–test rejection ratios ($R_N \approx$ 30%) during the 1-year testing period (Figures S16a and S16b). Models CRS-3/4 show better overall R_N values (15% and 14%, respectively) but underpredict ($\delta_1 < 0.025$) during the two weeks following the two largest events of the sequence (Figures S16c and S16d). We find that this is most probably due to the implementation of simple uniform slip models that project negative coseismic stress changes on the Mount Vettore-Mount Bove fault system after the Amatrice earthquake (see Figures S8c and S8d) and suppress the near-source expected rates after the Norcia mainshock with a heavy stress shadow (Figures S11s and S12d).

Although overforecasting in the first 24 hr following the Amatrice and Norcia earthquakes, CRS-5 and CRS-6 are the only two models that pass the test after the $M_w = 5.4$ Visso event and further reduce the overall R_N to 12% and 8.5%, respectively (Figures S16e and S16f).

We use the *S*–test joint log-likelihood scores (jLL_S) to evaluate the predictive power of the forecasts in space. We present the earthquake forecasts expressed by the joint log-likelihood over the entire testing region,

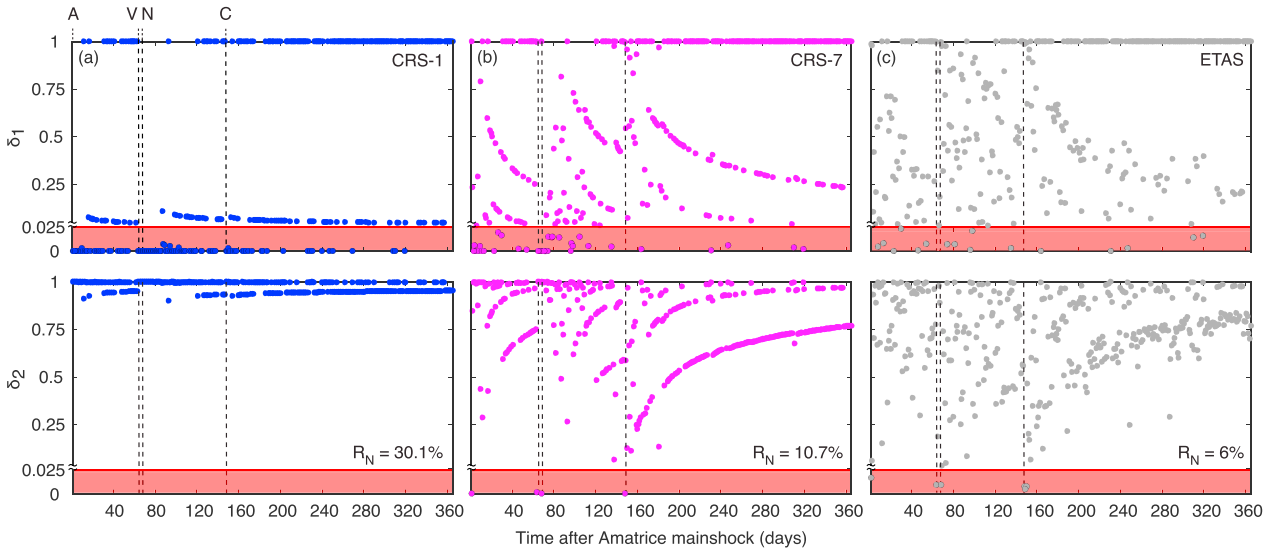

Figure 4. Incremental modified *N*–test scores. We show the δ_1 (top) and δ_2 (bottom) quantiles for CRS-1/7 and ETAS in the whole testing region, within 1-day intervals and for 1 year following the $M_w = 6.0$ Amatrice mainshock. The shaded areas under the red horizontal lines indicate the values for model rejection ($\delta_{1,2} < 0.025$); we apply a vertical exaggeration to these areas to better highlight the rejection values. Vertical lines mark mainshock occurrence (A = Amatrice, V = Visso, N = Norcia, C = Campotosto events). R_N = modified *N*–test rejection ratio. CRS = Coulomb rate-state; ETAS = Epidemic Type Aftershock Sequence.

where the expected rate at each spatial cell is multiplied by N_{obs}/N_{fore}, the ratio of forecasted to observed events in the entire region, so that the normalized forecast matches the observed number of events. Higher jLL_S values indicate a better model performance (Zechar et al., 2010). Results from the short-term performance test (Table 3) show that (1) the ETAS model outperforms all the stress-based counterparts in the first 24 hr after the Amatrice, Norcia, and Campotosto events; (2) the elaborate physics-based model CRS-7 presents the best jLLs score following the Visso earthquakes and it constantly achieves a better

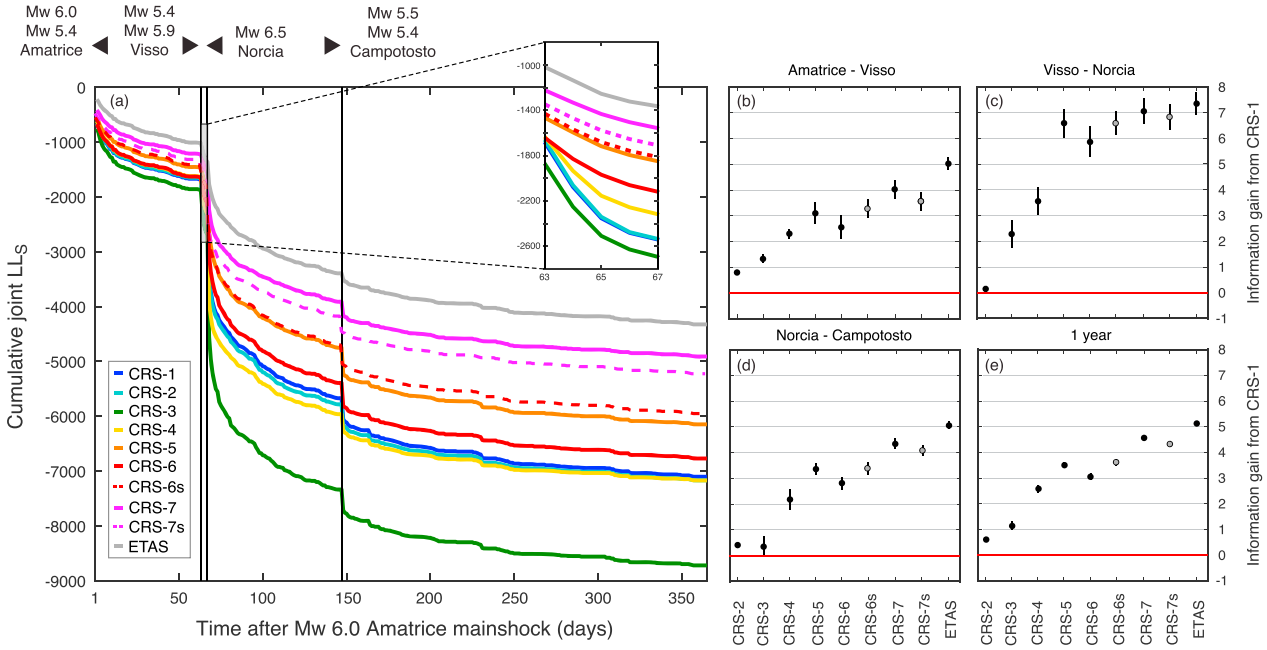

Figure 5. Cumulative spatial joint log-likelihood (jLL_S) versus time and *T*–test. (a) The scores are obtained by summing the *S*–test log-likelihoods (LL_S) of each spatial cell and 1-day time step. Vertical lines mark the occurrence of the largest magnitude events. (b–e) Average daily information gain per earthquake from the preliminary CRS-1 model for period-specific windows and for a 1-year forecast horizon. A forecast is significantly better than CRS-1 at 95% confidence if the error bars do not intersect the no gain level (red horizontal lines). CRS = Coulomb rate-state.

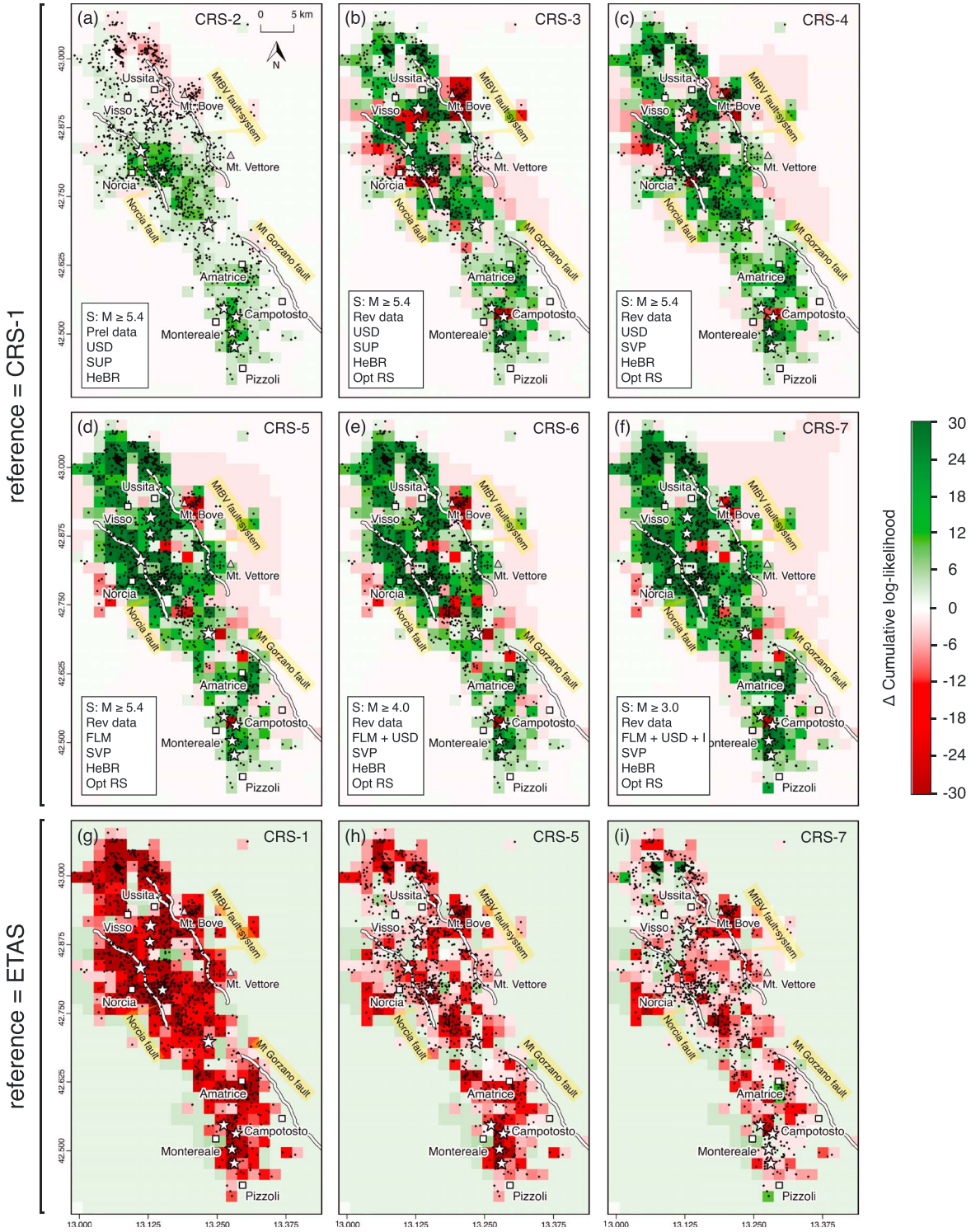

Figure 6. Spatially resolved cumulative log-likelihood differences (ΔLL) between pairs of models for a 1-year evaluation period. Maps show a subset of the testing region that includes the observed seismicity and the mainshock faults. As we do not need to isolate the spatial component of the forecasts, here we calculate the LL from unnormalized rates. LL values at each spatial bin are obtained summing over all time steps. Positive (green) values indicate improved performance with respect to the reference model specified to the left. Black dots represent the *M*3+ observations, while the stars indicate the *M*5+ events. Values are saturated at ±30 for visualization purposes. S = sources; Prel = preliminary; Rev = revised; USD = uniform slip distribution; FLM = finite-length rupture model; I = isotropic stress field; SUP = spatially uniform receiver planes; SVP = spatially variable planes; HoBR = spatially homogeneous background rate; HeBR = spatially heterogeneous background rate; Opt RS =optimized rate-state parameters; MtBV = Mount Bove-Vettore.

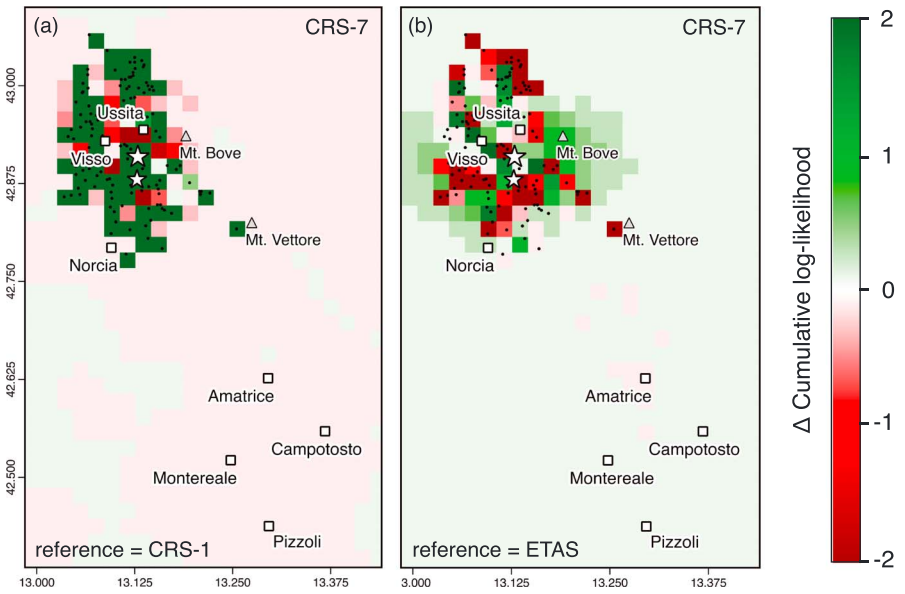

Figure 7. Maps of cumulative log-likelihood difference for the 4-day period between the Visso events and the Norcia mainshock between (a) CRS-7 and CRS-1, (b) CRS-7 and ETAS. Positive values indicate a better performance of CRS-7. Values are saturated at ±2 for clarity. CRS = Coulomb rate-state; ETAS = Epidemic Type Aftershock Sequence.

short-term performance than all the other CRS models; (3) CRS-5 is the second best CRS model in short-term windows, highlighting the importance of including spatially variable slip sources for physics-based modeling; the only exception is represented by the first 24 hr following the Campotosto seismicity where two out of the four $M5+$ earthquakes (namely the $M_w = 5.1$ and $M_w = 5.0$) lack a spatially variable rupture model, while for the $M_w = 5.4$ and $M_w = 5.5$ events only a preliminary slip model release is available.

In Figure 5a we compare the evolving spatial performance of the forecasts for the 1-year testing period after the $M_w = 6.0$ Amatrice occurrence, including the auxiliary models CRS-6s and CRS-7s that isolate the effect of secondary triggering from $M4+$ and $M3+$ sources, respectively. Results show that (1) CRS-7/7s, including secondary triggering effects from $M3+$ aftershocks, are the best physics-based models, particularly when rate-and-state parameters are optimized on the learning phase seismicity (CRS-7); (2) the ETAS model has the highest spatial joint log-likelihood values over the entire testing period ($jLL_S = -4326$), with a striking similar performance with CRS-7 in the 4 days between the Visso-Norcia earthquakes. While models CRS-1/2/4 perform similarly through the evaluation period, CRS-3 is characterized by the lowest scores because of the oversimplification of coseismic slip for the $M_w = 6.5$ Norcia mainshock. When we compare CRS-6 to CRS-6s, we note that the parameter optimization degrades CRS-6 performance after the Norcia mainshock. The long-term likelihood scores of CRS-6s are only slightly better than CRS-5, but a lower magnitude threshold for stress sources (CRS-7/7s) remarkably increases the short- and long-term spatial consistency of the models.

We estimate the information gains per earthquake using the preliminary CRS-1 as benchmark model ($G_{\text{CRS-1}}$). A positive G score indicates that a model is more informative than the selected benchmark. In Figures 5b–5e, we present the results in the form of average daily information gains per earthquake and their 95% confidence bounds from a paired Student's t-test (Rhoades et al., 2011) for CRS and ETAS forecasts. From the analysis of the G ranks, we note that CRS-7 is systematically the best performing among all CRS classes at all time periods (Figures 5b–5d and Table 3). The majority of CRS forecasts show positive information gains ($G_{\text{CRS-1}} \geq 2$), indicating a significant improvement on the preliminary CRS-1 model. For the critical ~4-day time window before the Norcia earthquake, CRS-7 and ETAS models present a comparable information gain ($\Delta G_{\text{CRS-1}} < 0.3$) with respect to the reference CRS-1, and CRS classes 5/6/7 show a considerably good performance ($G_{\text{CRS-1}} \approx 6$–7, Figure 5c). Especially in the first day after the Visso events, we find that CRS-7 outperforms the ETAS forecast (Table 3); the observed poorer performance of the

statistical model is most likely caused by the lack of $M3+$ precursory seismicity in the near-epicentral region of the first $M_w = 5.4$ Visso event (Chiaraluce et al., 2017). The 1-year performance evaluation reveals that the most advanced physics-based implementation, CRS-7, compares closely with the ETAS model ($\Delta G_{\text{CRS-1}} \approx 0.5$, Figure 5e). The trend of improving information gain from CRS-5 to CRS-6s/7s shows the importance behind considering stress changes from lower magnitude aftershocks. These results suggest that including SVP receivers, finite-length rupture models, and secondary triggering from $M3+$ sources in CRS models leads to a gradual information gain increase.

We assess the spatial component behind the models' performance by information gain maps for a 1-year time horizon (Figure 6), while period-specific maps are provided as supporting information (Figures S17–S19). We calculate the cumulative log-likelihood differences (ΔLL) at each grid point using CRS-1 (Figures 6a–6f) and ETAS (Figures 6g–6i) as reference models. In Figure 6, positive ΔLL values (green) highlight improved performance with respect to the reference model. Results show that (1) variable slip source representations lead to a sensible information gain increase in the Norcia-Visso near epicentral region (ΔLL up to 100) and on the Mount Bove-Vettore fault system ($\Delta\text{LL} > 60$; Figure 6d); (2) CRS-7 better captures the unfolding sequence, as considering the contribution of $M3+$ sources into the evolving stress field entails a net performance improvement especially at the extreme edges of the fault system ($\Delta\text{LL} > 250$ north of Visso and $\Delta\text{LL} > 50$ in the Montereale-Pizzoli region; Figure 6f); (3) updated source parameter description of mainshock faults through optimized rate-and-state parameters improves the forecast in many locations but reduces the likelihood in a broader aftershock region (Figure 6b), particularly at Mount Bove location ($\Delta\text{LL} < -100$); (4) model performance at the northwestern MtBV fault termination greatly benefits from the combined inclusion of SVP receiver planes (Figure 6c) and, to a greater extent, from finite-length slip models describing the Visso and Norcia coseismic slip (Figure 6d).

When we compare physics-based forecasts to ETAS, we observe that, although CRS models generally present a slightly positive information gain in the off-fault areas of the testing region ($\Delta\text{LL} \approx 0.3$), the log-likelihood differences between pairs of models are more marked where aftershocks were actually recorded. We find that ETAS widely outperforms its physics-based counterparts in the near-source regions (Figures 6g–6i), where slip models present large variability (Chiaraluce et al., 2017; Scognamiglio et al., 2016). However, we find that secondary triggering effects significantly increase the predictive power of CRS models that outperforms ETAS in the high clustering zone north of Visso-Ussita and in the Montereale-Pizzoli region (Figure 6i).

The model parameters that exert most influence on the predictive power of physics-based models are the spatially variable slip distributions and the inclusion of the secondary triggering effects from small magnitude earthquake ($M3+$). The most evolved CRS model captures the triggered seismicity at the fault edges, therefore improving the forecast in the case of the Campotosto cluster further south than Amatrice.

In Figure 7 we isolate the ΔLL in the critical period between the Visso and Norcia events, and we compare CRS-7 performance to CRS-1 and ETAS. We notice that CRS-7 is a significantly more informative model than CRS-1 in the time window that precedes the $M_w = 6.5$ mainshock (Figure 7a). Results also confirm the comparable performance between ETAS and CRS-7; the former better captures the near epicentral seismicity in the area of the Visso events, while the latter is better performing in locations of high aftershock production at fault extremities.

5. Conclusions

The complex 2016–2017 Amatrice-Visso-Norcia (AVN) earthquake sequence in the Central Apennines represents a unique opportunity to test earthquake triggering hypotheses that are the basis for physics-based forecast models. Here we focus on how input data quality and model parameters influence the predictive power of physics-based forecast models. We design a pseudoprospective experiment for the first year of the AVN sequence. We implement a benchmark statistical ETAS model and seven classes of physics-based forecasts with gradually increasing level of input data quality and complexity. We then evaluate the absolute and relative model performance by means of the N (number), S (space) and T–test metrics that are currently implemented within the CSEP initiative.

We find that the representation of the crustal and stress-field heterogeneity increases the information gains of physics-based models. The CRS model that predicts closely and in cases outperforms the statistical ETAS

model combines best-available source models, spatially varying receivers informed by preexisting focal mechanisms and geological maps, and most importantly includes secondary triggering effects from $M3+$ events that enable us to describe the evolving coseismic stress field in greater detail. After the $M_w = 5.4$ Visso event, the elaborate CRS-7 model outperforms ETAS, while in the critical 4-day time window before the $M_w = 6.5$ Norcia mainshock and within the first year of the sequence, CRS-7 is as informative as ETAS.

The results support previous findings that in the near-source area, defined by the surface fault projection, ETAS models present higher predictive power (e.g., Segou et al., 2013), although the lack of precursory seismicity (e.g., before the Visso events) locally hampers their short-term performance. On the other hand, triggered seismicity at intermediate off-fault distances due to static stresses within an evolving sequence is better described by physics-based approaches. The preliminary unrevised source parameters and empirical source models released within a few minutes to hours following large events result in poor performance of stress-based models that by definition require a more demanding physical parametrization than their empirical counterparts. The comparative evaluation of the spatial consistency of CRS models suggests that the over-simplified uniform slip models implemented immediately after a large earthquake cannot adequately reproduce the early aftershock spatial distribution (< 1 day).

The Italian experiment reveals that in order for stress-based forecasts to reach and outperform ETAS models we need to use best-available data within heterogeneous fault and source representations. Although high-quality relocated catalogs are not yet readily available during the early stages of a seismic crisis, both CRS and ETAS models in our study will improve their predictive skills with enhanced detection and event characterization techniques.

Acknowledgments
The authors thank Andrea Llenos and an anonymous reviewer for their constructive reviews. We also thank Tom Parsons (USGS) for his interesting comments that helped us improving the manuscript. We are grateful to Elisa Tinti (INGV-Rome) for providing the rupture models. S. Mancini thanks Stefanie Seif for providing the code for estimating ETAS parameters and performing simulations. The code *CRS* (Cattania & Khalid, 2016) can be found at https://github.com/camcat/crs. The earthquake catalog for the AVN sequence with event locations and magnitude chronologies was acquired through access to http://cnt.rm.ingv.it, while the moment tensor solutions from http://cnt.rm.ingv.it/tdmt. The Italian CMT data set is available at http://rcmt2.bo.ingv.it/Italydataset.html. The Database of Individual Seismogenic Sources for Italy and surrounding areas is accessible at http://diss.rm.ingv.it/diss. S. Mancini is supported by a NERC GW4+ Doctoral Training Partnership studentship from the Natural Environment Research Council (NE/L002434/1) and by the BGS University Funding Initiative Ph. D. studentship (S350). M. Werner is supported by the Southern California Earthquake Center (contribution 9015). SCEC is funded by NSF Cooperative Agreement EAR-1600087 and USGS Cooperative Agreement G17AC00047. The authors thank the participants to the NERC-NSF supported project *The Central Apennines Earthquake Sequence under a New Microscope* (NE/R0000794/1) for facilitating fruitful discussion about the state-of-the-art knowledge on the seismic sequence.

References

Amato, A., Azzara, R., Chiarabba, C., Cimini, G. B., Cocco, M., Di Bona, M., et al. (1998). The 1997 Umbria-Marche, Italy, earthquake sequence: A first look at the main shocks and aftershocks. *Geophysical Research Letters*, *25*(15), 2861–2864. https://doi.org/10.1029/98GL51842

Basili, R., Valensise, G., Vannoli, P., Burrato, P., Fracassi, U., Mariano, S., et al. (2008). The Database of Individual Seismogenic Sources (DISS), version 3: Summarizing 20 years of research on Italy's earthquake geology. *Tectonophysics*, *453*(1-4), 20–43. https://doi.org/10.1016/j.tecto.2007.04.014

Boschi, E., Guidoboni, E., Ferrari, G., Mariotti, D., Valensise, G., & Gasperini, P. (Eds) (2000). Catalogue of strong Italian earthquakes from 461 B.C. to 1997. *Annales de Geophysique*, *43*(4), 609–868. https://doi.org/10.4401/ag-3668

Catalli, F., Cocco, M., Console, R., & Chiaraluce, L. (2008). Modeling seismicity changes during the 1997 Umbria-Marche sequence (central Italy) through a rate- and state-dependent model. *Journal of Geophysical Research*, *113*, B11301. https://doi.org/10.1029/2007JB005356

Cattania, C., & Khalid, F. (2016). A parallel code to calculate rate-state seismicity evolution induced by time dependent, heterogeneous Coulomb stress changes. *Computers & Geosciences*, *94*, 48–55. https://doi.org/10.1016/j.cageo.2016.06.007

Cattania, C., Werner, M. J., Marzocchi, W., Hainzl, S., Rhoades, D., Gerstenberger, M., et al. (2018). The forecasting skill of physics-based seismicity models during the 2010–2012 Canterbury, New Zealand, earthquake sequence. *Seismological Research Letters*, *89*(4), 1238–1250. https://doi.org/10.1785/0220180033

Chen, K. H., Bürgmann, R., & Nadeau, R. M. M. (2013). Do earthquakes talk to each other? Triggering and interaction of repeating sequences at Parkfield. *Journal of Geophysical Research: Solid Earth*, *118*, 165–182. https://doi.org/10.1029/2012JB009486

Chiarabba, C., Amato, A., Anselmi, M., Baccheschi, P., Bianchi, I., Cattaneo, M., et al. (2009). The 2009 L'Aquila (central Italy) M_w6.3 earthquake: Main shock and aftershocks. *Geophysical Research Letters*, *36*, L18308. https://doi.org/10.1029/2009GL039627

Chiaraluce, L., Amato, A., Cocco, M., Chiarabba, C., Selvaggi, G., Di Bona, M., et al. (2004). Complex normal faulting in the Apennines thrust-and-fold belt: The 1997 seismic sequence in central Italy. *Bulletin of the Seismological Society of America*, *94*(1), 99–116. https://doi.org/10.1785/0120020052

Chiaraluce, L., Di Stefano, R., Tinti, E., Scognamiglio, L., Michele, M., Casarotti, E., et al. (2017). The 2016 Central Italy seismic sequence: A first look at the mainshocks, aftershocks, and source models. *Seismological Research Letters*, *88*(3), 757–771. https://doi.org/10.1785/0220160221

Chiaraluce, L., Ellsworth, W. L., Chiarabba, C., & Cocco, M. (2003). Imaging the complexity of an active normal fault system: The 1997 Colfiorito (central Italy) case study. *Journal of Geophysical Research*, *108*(B6). https://doi.org/10.1029/2002JB002166

Chiaraluce, L., Valoroso, L., Piccinini, D., Di Stefano, R., & De Gori, P. (2011). The anatomy of the 2009 L'Aquila normal fault system (central Italy) imaged by high resolution foreshock and aftershock locations. *Journal of Geophysical Research*, *116*(B12). https://doi.org/10.1029/2011JB008352

Cocco, M., Hainzl, S., Catalli, F., Enescu, B., Lombardi, A. M., & Woessner, J. (2010). Sensitivity study of forecasted aftershock seismicity based on Coulomb stress calculation and rate- and state-dependent frictional response. *Journal of Geophysical Research*, *115*(B5), B05307. https://doi.org/10.1029/2009JB006838

Console, R., Murru, M., & Catalli, F. (2006). Physical and stochastic models of earthquake clustering. *Tectonophysics*, *417*(1-2), 141–153. https://doi.org/10.1016/j.tecto.2005.05.052

Dieterich, J. H. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research*, *99*(B2), 2601–2618. https://doi.org/10.1029/93JB02581

DISS Working Group (2018). Database of Individual Seismogenic Sources (DISS), Version 3.2.1: A compilation of potential sources for earthquakes larger than *M* 5.5 in Italy and surrounding areas. http://diss.rm.ingv.it/diss/, Istituto Nazionale di Geofisica e Vulcanologia. https://doi.org/10.6092/INGV.IT-DISS3.2.1.

EMERGEO Working Group (2016). The 24 August 2016 Amatrice earthquake: Coseismic effects. https://doi.org/10.5281/zenodo.61568.

Field, E. H., Arrowsmith, R. J., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., et al. (2014). Uniform California Earthquake Rupture Forecast, version 3 (UCERF3)—The time-independent model. *Bulletin of the Seismological Society of America*, *104*(3), 1122–1180. https://doi.org/10.1785/0120130164

Hanks, T., & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research*, *84*(B5), 2348–2350. https://doi.org/10.1029/JB084iB05p02348

Harris, R. A., & Simpson, R. W. (1992). Changes in static stress on southern California faults after the 1992 Landers earthquake. *Nature*, *360*(6401), 251–254. https://doi.org/10.1038/360251a0

Helmstetter, A., Kagan, Y., & Jackson, D. D. (2005). Importance of small earthquakes for stress transfers and earthquake triggering. *Journal of Geophysical Research*, *110*, B05S08. https://doi.org/10.1029/2004JB003286

Helmstetter, A., Kagan, Y., & Jackson, D. D. (2006). Comparison of short-term and time-independent earthquake forecast models for Southern California. *Bulletin of the Seismological Society of America*, *96*(1), 90–106. https://doi.org/10.1785/0120050067

Helmstetter, A., Kagan, Y., & Jackson, D. D. (2007). High-resolution time-independent grid-based forecast for $M >= 5$ earthquakes in California. *Seismological Research Letters*, *78*(1), 78–86. https://doi.org/10.1785/gssrl.78.1.78

Jordan, T. H. (2006). Earthquake predictability, brick by brick. *Seismological Research Letters*, *77*(1), 3–6. https://doi.org/10.1785/gssrl.77.1.3

Jordan, T. H., Chen, Y., & Main, I. (2011). Operational earthquake forecasting: State of knowledge and guidelines for utilization. *Annales de Geophysique*, *54*(4), 316–391. https://doi.org/10.4401/ag-5350

Jordan, T. H., Marzocchi, W., Michael, A. J., & Gerstenberger, M. C. (2014). Operational earthquake forecasting can enhance earthquake preparedness. *Seismological Research Letters*, *85*(5), 955–959. https://doi.org/10.1785/0220140143

King, G. C. P., Stein, R. S., & Lin, J. (1994). Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America*, *84*(3), 935–953.

Marsan, D. (2005). The role of small earthquakes in redistributing crustal elastic stress. *Geophysical Journal International*, *163*(1), 141–151. https://doi.org/10.1111/j.1365-246X.2005.02700.x

Marzocchi, W., & Lombardi, A. M. (2009). Real-time forecasting following a damaging earthquake. *Geophysical Research Letters*, *36*, L21302. https://doi.org/10.1029/2009GL040233

Marzocchi, W., Murru, M., Lombardi, A. M., Falcone, G., & Console, R. (2012). Daily earthquake forecasts during the May-June 2012 Emilia earthquake sequence (northern Italy). *Annales de Geophysique*, *55*(4), 561–567. https://doi.org/10.4401/ag-6161

Marzocchi, W., Taroni, M., & Falcone, G. (2017). Earthquake forecasting during the complex Amatrice-Norcia seismic sequence. *Science Advances*, *3*(9), e1701239. https://doi.org/10.1126/sciadv.1701239

Marzocchi, W., Zechar, J. D., & Jordan, T. H. (2012). Bayesian forecast evaluation and ensemble earthquake forecasting. *Bulletin of the Seismological Society of America*, *102*(6), 2574–2584. https://doi.org/10.1785/0120110327

McCloskey, J., Nalbant, S. S., Steacy, S., Nostro, C., Scotti, O., & Baumont, D. (2003). Structural constraints on the spatial distribution of aftershocks. *Geophysical Research Letters*, *30*(12), 1610. https://doi.org/10.1029/2003GL017225

Michael, A. J., & Werner, M. J. (2018). Preface to the focus section on the collaboratory for the Study of Earthquake Predictability (CSEP): New results and future directions. *Seismological Research Letters*, *89*(4), 1226–1228. https://doi.org/10.1785/0220180161

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, *83*(401), 9–27. https://doi.org/10.1080/01621459.1988.10478560

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, *50*(2), 379–402. https://doi.org/10.1023/A:1003403601725

Ogata, Y., & Zhuang, J. (2006). Space-time ETAS models and an improved extension. *Tectonophysics*, *413*(1–2), 13–23. https://doi.org/10.1016/j.tecto.2005.10.016

Omi, T., Ogata, Y., Shiomi, K., Enescu, B., Sawazaki, K., & Aihara, K. (2016). Automatic aftershock forecasting: A test using real-time seismicity data in Japan. *Bulletin of the Seismological Society of America*, *106*(6), 2450–2458. https://doi.org/10.1785/0120160100

Parsons, T., Ogata, Y., Zhuang, Z., & Geist, E. L. (2012). Evaluation of static stress change forecasting with prospective and blind tests. *Geophysical Journal International*, *188*(3), 1425–1440. https://doi.org/10.1111/j.1365-246X.2011.05343.x

Parsons, T., Segou, M., Sevilgen, V., Milner, K., Field, E., Toda, S., & Stein, R. S. (2014). Stress based aftershock forecasts made within 24 h postmain shock: Expected north San Francisco Bay area seismicity changes after the 2014 $M = 6.0$ west Napa earthquake. *Geophysical Research Letters*, *41*, 8792–8799. https://doi.org/10.1002/2014GL062379

Rhoades, D. A., Schorlemmer, D., Gerstenberger, M. C., Christophersen, A., Zechar, J. D., & Imoto, M. (2011). Efficient testing of earthquake forecasting models. *Acta Geophysica*, *59*(4), 728–747. https://doi.org/10.2478/s11600-011-0013-5

Rice, J. R. (1992). Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault. In B. Evans, & T. Wong (Eds.), *Fault mechanics and transport properties of rocks; a festschrift in honour of W. F. Brace*, edited by, (Chapter 20, p. 475, 503). San Diego, CA, USA.: Academic Press.

Rovida, A., Locati, M., Camassi, R., Lolli, B., & Gasperini, P. (Eds) (2016). *CPTI15, the 2015 Version of the Parametric Catalogue of Italian Earthquakes*. Rome, Italy: Istituto Nazionale di Geofisica e Vulcanologia. https://doi.org/10.6092/INGV.IT-CPTI15

Schorlemmer, D., Gerstenberger, M. C., Wiemer, S., Jackson, D. D., & Rhoades, D. A. (2007). Earthquake likelihood model testing. *Seismological Research Letters*, *78*(1), 17–29. https://doi.org/doi:10.1785/gssrl.78.1.17

Schorlemmer, D., Mele, F., & Marzocchi, W. (2010). A completeness analysis of the National Seismic Network of Italy. *Journal of Geophysical Research*, *115*(B4), B04308. https://doi.org/10.1029/2008JB006097

Scognamiglio, L., Tinti, E., & Quintiliani, M. (2016). The first month of the 2016 Amatrice seismic sequence: Fast determination of time domain moment tensors and finite fault model analysis of the ML 5.4 aftershock. *Annales de Geophysique*, *59*, 1–10. https://doi.org/10.4401/ag-7246

Segou, M. (2016). Physics-based and statistical earthquake forecasting in a continental rift zone: The case study of Corinth Gulf (Greece). *Geophysical Journal International*, *204*(1), 591–605. https://doi.org/10.1093/gji/ggv467

Segou, M., & Parsons, T. (2016). Prospective earthquake forecasts at the Himalayan Front after the 25 April 2015 M 7.8 Gorkha mainshock. *Seismological Research Letters*, *87*(4), 816–825. https://doi.org/10.1785/0220150195

Segou, M., Parsons, T., & Ellsworth, W. (2013). Comparative evaluation of physics-based and statistical forecasts in Northern California. *Journal of Geophysical Research: Solid Earth*, *118*, 6219–6240. https://doi.org/10.1002/2013JB010313

Seif, S., Mignan, A., Zechar, J. D., Werner, M. J., & Wiemer, S. (2016). Estimating ETAS: The effects of truncation, missing data, and model assumptions. *Journal of Geophysical Research: Solid Earth*, *122*, 449–469. https://doi.org/10.1002/2016JB012809

Serpelloni, E., Anzidei, M., Baldi, P., Casula, G., & Galvani, A. (2005). Crustal velocity and strain-rate fields in Italy and surrounding regions: New results from the analysis of permanent and non-permanent GPS networks. *Geophysical Journal International*, *161*(3), 861–880. https://doi.org/10.1111/j.1365-246X.2005.02618.x

Stein, S., & Liu, M. (2009). Long aftershock sequences within continents and implications for earthquake hazard assessment. *Nature Letters*, *462*(7269), 87–89. https://doi.org/10.1038/nature08502

Tinti, E., Scognamiglio, L., Michelini, A., & Cocco, M. (2016). Slip heterogeneity and directivity of the ML 6.0, 2016, Amatrice earthquake estimated with rapid finite-fault inversion. *Geophysical Research Letters*, *43*, 10,745–10,752. https://doi.org/10.1002/2016GL071263

Toda, S., & Enescu, B. (2011). Rate/state Coulomb stress transfer model for the CSEP Japan seismicity forecast. *Earth, Planets and Space*, *63*(3), 171–185. https://doi.org/10.5047/eps.2011.01.004

Toda, S., Stein, R. S., Reasenberg, P. A., & Dieterich, J. H. (1998). Stress transferred by the 1995 $M_w = 6.9$ Kobe, Japan, shock: Effect in aftershocks and future earthquake probabilities. *Journal of Geophysical Research*, *103*(B10), 543–565.

Toda, S., Stein, R. S., Richards-Dinger, K., & Bozkurt, B. S. (2005). Forecasting the evolution of seismicity in southern California: Animations built on earthquake stress transfer. *Journal of Geophysical Research*, *110*(B5), B05S16. https://doi.org/10.1029/2004JB003415

Utsu, T. (1961). A statistical study on the occurrence of aftershocks. *Geophysical Magazine*, *30*, 521–605.

Wells, D. L., & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America*, *84*, 974–1002.

Werner, M. J., Helmstetter, A., Jackson, D. D., & Kagan, Y. Y. (2011). High-resolution long-term and short-term earthquake forecasts for California. *Bulletin of the Seismological Society of America*, *101*(4), 1630–1648. https://doi.org/doi:10.1785/0120090340

Woessner, J., Hainzl, S., Marzocchi, W., Werner, M. J., Lombardi, A. M., Catalli, F., et al. (2011). A retrospective comparative forecast test on the 1992 Landers sequence. *Journal of Geophysical Research*, *116*(B5), B05305. https://doi.org/10.1029/2010JB007846

Woessner, J., Laurentiu, D., Giardini, D., Crowley, H., Cotton, F., Grunthal, G., et al. (2015). The 2013 European seismic hazard model: Key components and results. *Bulletin of Earthquake Engineering*, *13*(12), 3553–3596. https://doi.org/10.1007/s10518-015-9795-1

Zechar, J. D., Gerstenberger, M. C., & Rhoades, D. A. (2010). Likelihood based tests for evaluating space-rate-magnitude earthquake forecasts. *Bulletin of the Seismological Society of America*, *100*(3), 1184–1195. https://doi.org/10.1785/0120090192

Zhuang, J., Harte, D., Werner, M. J., Hainzl, S., & Zhou, S. (2012). Basic models of seismicity: Temporal models. In *Community Online Resource for Statistical Seismicity Analysis*. Chapter (Theme) 5, 1–45, CORSSA. https://doi.org/10.5078/corssa-79905851

Zhuang, J., Ogata, Y., & Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association*, *97*(458), 369–380. https://doi.org/10.1198/016214502760046925

References From the Supporting Information

Buttinelli, M., Pezzo, G., Valoroso, L., De Gori, P., & Chiarabba, C. (2018). Tectonics inversions, fault segmentation, and triggering mechanisms in the central Apennines normal fault system: Insights from high-resolution velocity models. *Tectonics*, *37*(11), 4135–4149. https://doi.org/10.1029/2018TC005053

Cao, A. M., & Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern Japan island arc. *Geophysical Research Letters*, *29*(9), 1334. https://doi.org/10.1029/2001GL013775

Cattania, C., Hainzl, S., Wang, L., Roth, F., & Enescu, B. (2014). Propagation of Coulomb stress uncertainties in physics-based aftershock models. *Journal of Geophysical Research: Solid Earth*, *119*, 7846–7864. https://doi.org/10.1002/2014JB011183

Chiarabba, C., & De Gori, P. (2016). The seismogenic thickness in Italy: Constraints on potential magnitude and seismic hazard. *Terra Nova*, *28*(6), 402–408. https://doi.org/10.1111/ter.12233

Chiarabba, C., De Gori, P., Cattaneo, M., Spallarossa, D., & Segou, M. (2018). Faults geometry and the role of fluids in the 2016–2017 Central Italy seismic sequence. *Geophysical Research Letters*, *45*(14), 6963–6971. https://doi.org/10.1029/2018GL077485

Chiarabba, C., Jovane, L., & Di Stefano, R. (2005). A new view of Italian seismicity using 20 years of instrumental recordings. *Tectonophysics*, *395*(3-4), 251–268. https://doi.org/10.1016/j.tecto.2004.09.013

Cocco, M., & Rice, J. R. (2002). Pore pressure and poroelasticity effects in Coulomb stress analysis of earthquake interactions. *Journal of Geophysical Research*, *107*(B2), 2030. https://doi.org/10.1029/2000JB000138

Hainzl, S., Enescu, B., Cocco, M., Woessner, J., Catalli, F., Wang, R., & Roth, F. (2009). Aftershock modeling based on uncertain stress calculations. *Journal of Geophysical Research*, *114*(B5), B05309. https://doi.org/10.1029/2008JB006011

Malagnini, L., Lucente, F. P., De Gori, P., Akinci, A., & Munafo, I. (2012). Control of pore fluid pressure diffusion on fault failure mode: Insights from the 2009 L'Aquila seismic sequence. *Journal of Geophysical Research*, *117*(B5), B05302. https://doi.org/10.1029/2011JB008911

Rong, Y., Bird, P., & Jackson, D. D. (2016). Earthquake potential and magnitude limits inferred from a geodetic strain-rate model for southern Europe. *Geophysical Journal International*, *205*(1), 509–522. https://doi.org/10.1093/gji/ggw018

Wiemer, S. (2001). A software package to analyze seismicity: ZMAP. *Seismological Research Letters*, *72*(3), 373–382. https://doi.org/10.1785/gssrl.72.3.373

Wiemer, S., & Wyss, M. (2000). Minimum magnitude of complete reporting in earthquake catalogs: Examples from Alaska, the western United States, and Japan. *Bulletin of the Seismological Society of America*, *90*(4), 859–869. https://doi.org/10.1785/0119990114

Woessner, J., Jónsson, S., Sudhaus, H., & Baumann, C. (2012). Reliability of Coulomb stress changes inferred from correlated uncertainties of finite-fault source models. *Journal of Geophysical Research*, *117*(B7), B07303. https://doi.org/10.1029/2011JB009121

Woessner, J., & Wiemer, S. (2005). Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America*, *95*(2), 684–698. https://doi.org/10.1785/0120040007

Wyss, M., Hasegawa, A., Wiemer, S., & Umino, N. (1999). Quantitative mapping of precursory seismic quiescence before the 1989, *M*7.1 Osanriku earthquake, Japan. *Annali Di Geofsica*, *42*, 851–869. https://doi.org/10.4401/ag-3765

Zhuang, J., & Touati, S. (2015). Stochastic simulation of earthquake catalogs. In *Community Online Resource for Statistical Seismicity Analysis*. Chapter (Theme) 5, 1–34, CORSSA. https://doi.org/10.5078/corssa-43806322